

Monte Carlo Simulation for the EIC: Status and Experimental Requirements

MC4EIC

CTEQ-EICUG workshop on MC event simulation for the EIC

November 18-19, 2021 <https://indico.bnl.gov/event/13298/>

Markus Diefenthaler



Measurements and simulations in experimental Nuclear Physics

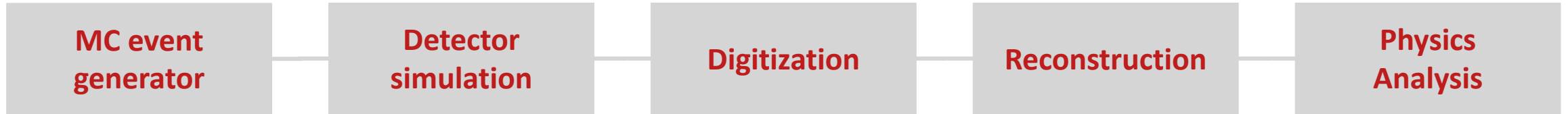
Measurement



Simulation



Role of simulations in experimental Nuclear Physics



Design Experiments Design and develop detectors and large-scale detector systems based on key measurements / physics reach and background estimates. Optimize the design.

Analysis Develop and verify analysis methods and tools: Does the analysis tool or method give the correct result? Estimate systematic uncertainties.

Verify Measurements Detailed simulations essential for commissioning experiments and verify analyses.

Simulation Software for the EIC

Simulation of physics processes

Monte Carlo Event Generators

Simulation of detector responses

Fast simulations

Full simulations

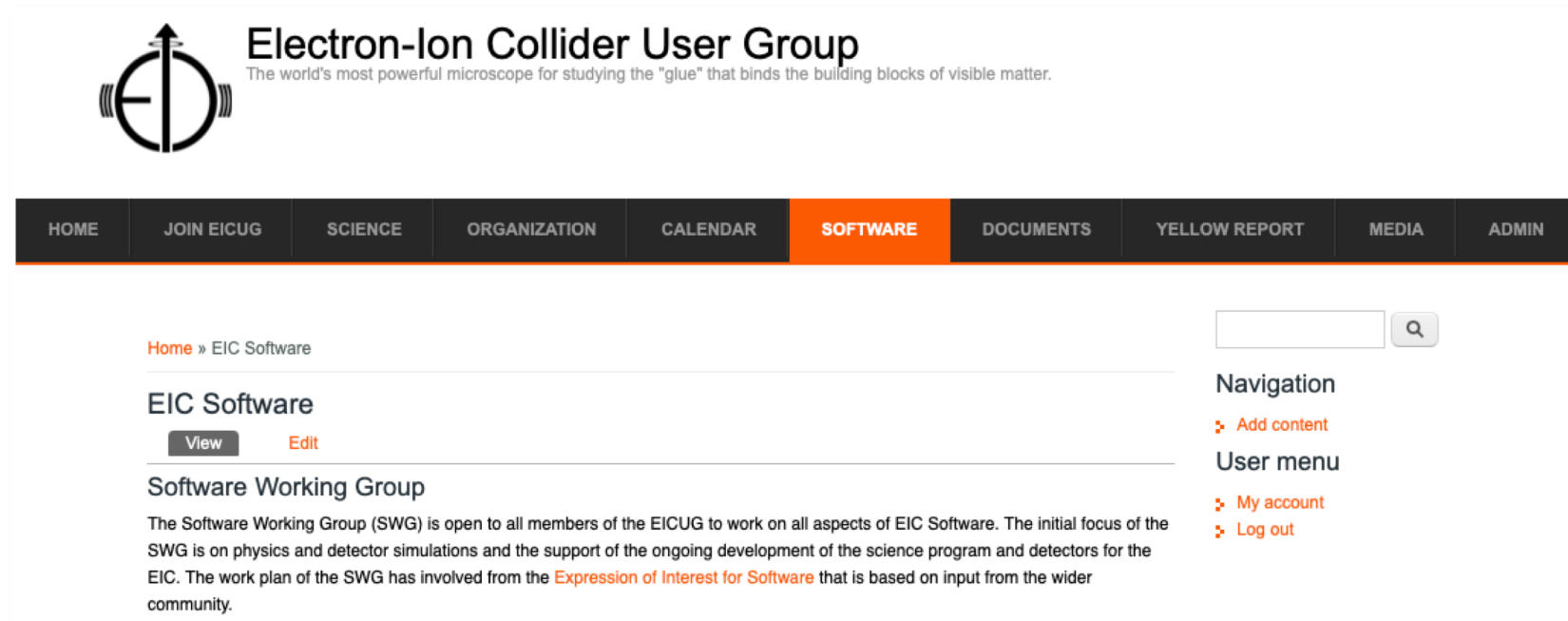
Analysis of simulated data

Reconstruction

EIC R&D For Software & Computing

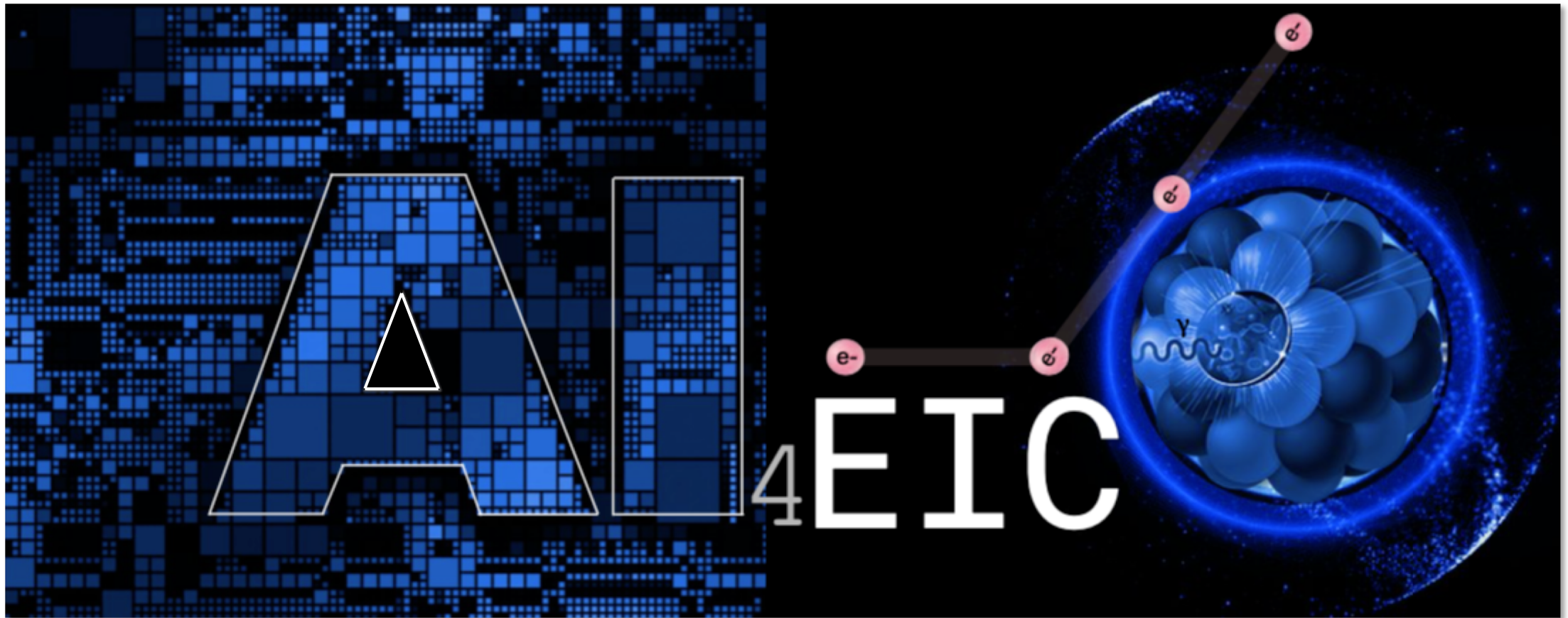
EIC Software & Computing is in a very early life stage:

- The current focus is supporting detector design.
- **Software Working Group** (SWG) within the EIC User Group works with the community and proto-collaborations to address software needs and evolving R&D.
- Legacy codes and frameworks are in use.
- Distributed Computing approach to supply resources for physics detector studies.
- At the pre-requirements stage for production computing and software activities.



AI/ML for EIC

AI/ML already has an important presence in EIC, with one of the proto-collaborations (ECCE) applying it to detector design optimization, as well as applications such as streaming DAQ, and a **new** AI Working Group as part of SWG to explore and develop AI/ML's potential.



Driven by community need, 'bottom up'



ELECTRON ION COLLIDER USER GROUP STATE OF SOFTWARE SURVEY

The Software Working Group collected information on the community's specific software tools and practices during the Yellow Report Initiative.

Q7. Do you have any comments on your current experience with EIC Software?

**Common message:
Priorities for consolidating around
common software are in MC
generators and detector simulation**

EIC Common Software: The Software EOI (and 'living' planning document)

Call for Expressions of Interest for Potential Cooperation on the EIC Experimental Program

Brookhaven National Laboratory
Expression of Interest (EOI) for
Electron-Ion Collider (EIC). The
integrated in the interaction region

Expression of Interest (EOI) for Software

Please indicate the name of the contact person for this submission:

Conveners of the Software Working Group:

- A. Bressan, M. Dieffenthaler, and T. Wenaus
- eicug-software-conveners@eicug.org

Please indicate all institutions collectively involved in this submission of interest:

ANL	Argonne National Laboratory
BNL	Brookhaven National Laboratory
CEA/Irfu	IRFU at CEA /Saclay institute
EIC-India	Alkal University, Central University of Karnataka, DAV College Chandigarh, Goa University, Indian Institute of Technology Bombay, Indian Institute of Technology Delhi, Indian Institute of Technology Indore, Indian Institute of Technology Patna, Indian Institute of Technology Madras, Malaviya National Institute of Technology Jaipur, Panjab University, Ramkrishna Mission Residential College Kolkata
IMP-CAS	Institute of Modern Physics - Chinese Academy of Sciences
INFN	Istituto Nazionale di Fisica Nucleare
JLab	Thomas Jefferson National Accelerator Facility
LANL	Los Alamos National Laboratory
LBNL and UC Berkeley	Lawrence Berkeley National Laboratory and University of California, Berkeley
NCBJ	National Centre for Nuclear Research
OhioU	Ohio University
ORNL	Oak Ridge National Laboratory
SBU	Stony Brook University
SLAC	SLAC National Accelerator Laboratory
SU	Shandong University

- **Software Tools for Simulations and Reconstruction**
 - Monte Carlo Event Generators
 - Detector Simulations
 - Reconstruction
 - Validation
- **Middleware and Preservation**
 - Workflows
 - Data and Analysis Preservation
- **Interaction with the Software Tools**
 - Explore User-Centered Design
 - Discoverable Software cvmfs/spack
 - Data Model
- **Future Technologies**
 - Heterogeneous computing
 - New languages and tools
 - Collaborative software

<https://indico.bnl.gov/event/8552/contributions/43221/>

MCEG Distribution for the EIC



Kolja Kauder (BNL)

EIC community has been organized around its MCEGs needs already for several years:

- PYTHIA6 (modified)
- BeAGLE
- DJANGO
- elSpectro
- eSTARlight
- MILOU
- PEPSI
- RAPGAP
- Sartre
- TOPEG (Orsay Perugia)

Maintained on CVMFS and used for a plethora of EIC studies. Established HepMC2/3 as standard in the wider EIC community (thanks to Andrii Verbytskyi (MPP) for support).

Discussion of Event Generation and Simulation Needs

Simulation of physics processes

Monte Carlo Event Generators

Simulation of detector responses

Fast simulations

Full simulations

Analysis of simulated data

Reconstruction

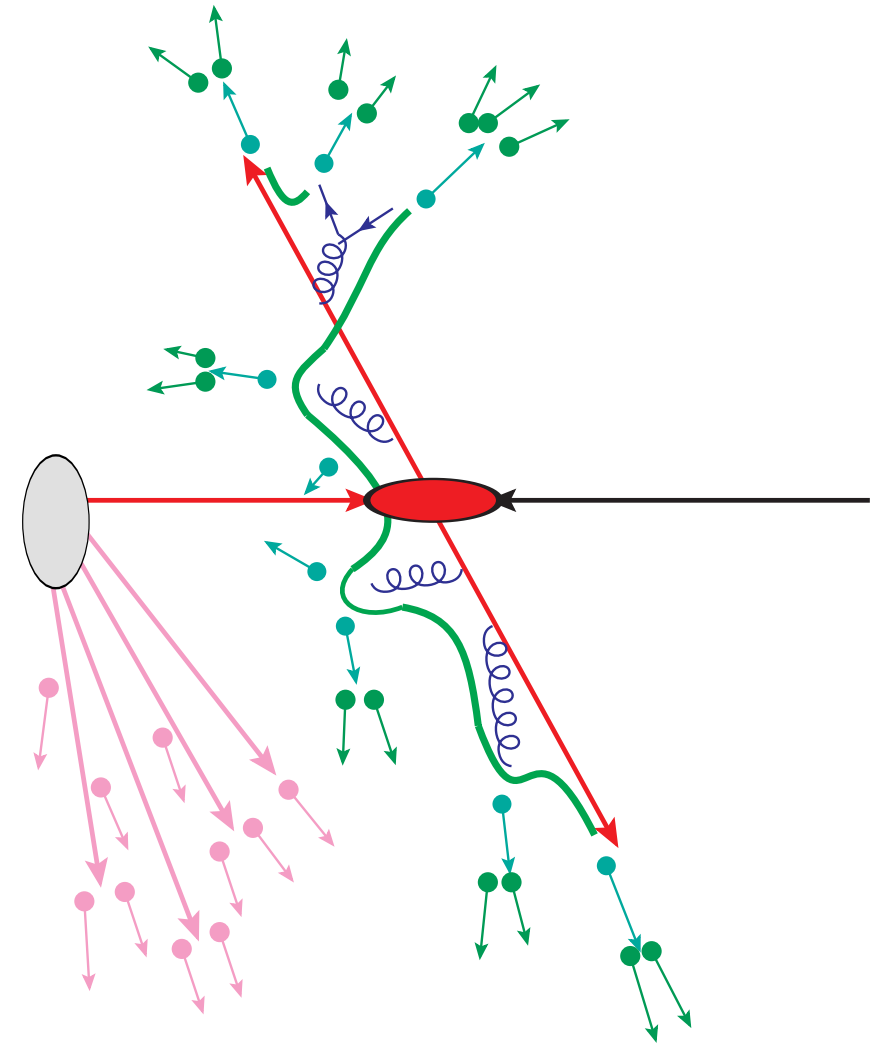
Monte Carlo Event Generator

MCEG

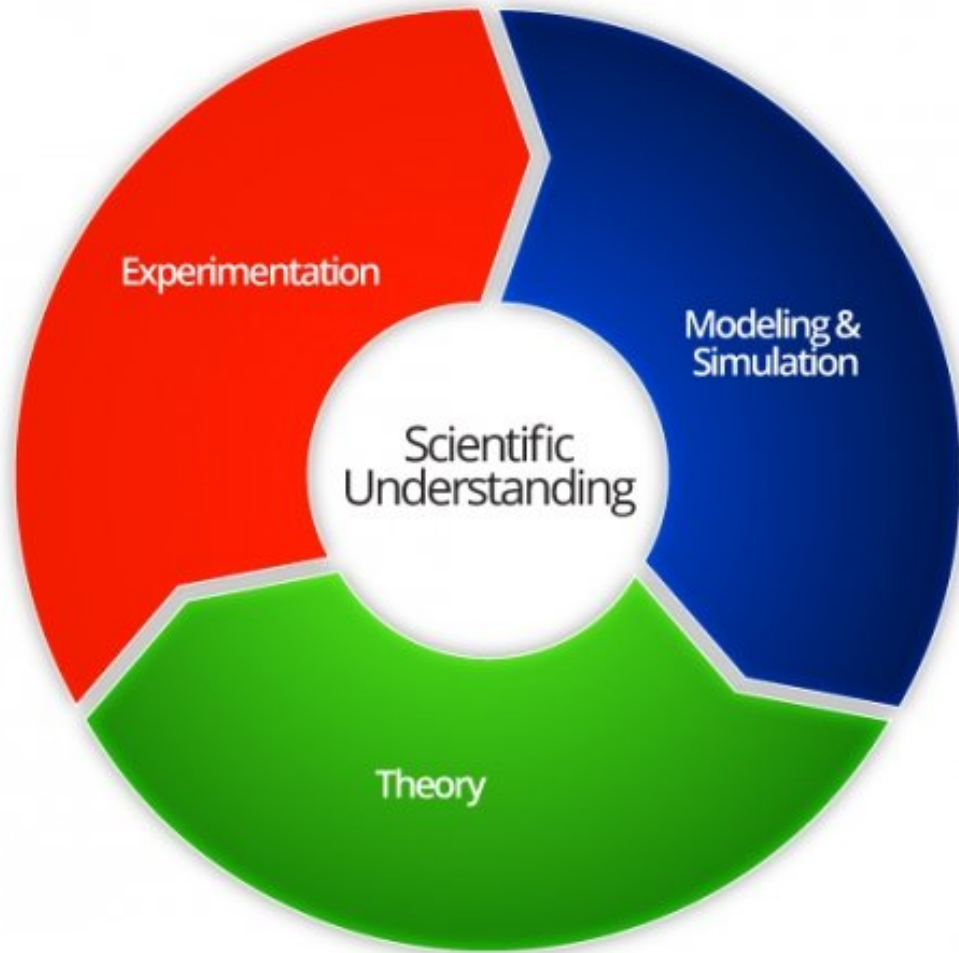
- faithful representation of QCD dynamics
- based on QCD factorization and evolution equations

MCEG algorithm

1. Generate kinematics according to fixed-order matrix elements and a PDF. **Session on "The Role of PDFs"**
2. QCD Evolution via parton shower model (resummation of soft gluons and parton-parton scatterings). **Session on "The Role of Parton Showers"**
3. Hadronize all outgoing partons including the remnants according to a model. **Session on "The Role of Hadronization"**
4. Decay unstable hadrons.



Event Generators for the EIC



Monte Carlo Simulation of

- electron-proton (ep) collisions,
- electron-ion (eA) collisions, both light and heavy ions,
- including higher order QED and QCD effects,
- including a plethora of spin-dependent effects.

Common challenges, e.g. with HL-LHC: **High-precision QCD measurements require high-precision simulations.**

Unique challenges MCEGs for electron-**ion** collisions and **spin-dependent** measurements, including novel QCD phenomena (e.g., GPDs or TMDs).

Start building a MCEG community for the EIC

The poster for POETIC 8 features a dark blue background with a silhouette of a city skyline and a large, colorful, abstract circular pattern resembling a particle detector or a complex network. The text is white and yellow.

P O E T I C 8
8th International Conference on Physics Opportunities at an Electron-Ion Collider
19-23 March 2018, University of Regensburg

Local Organizing Committee:
Gunnar Bali
Vladimir Braun
Falk Bruckmann
Sara Collins
Andreas Schäfer (chair)
Stefan Solbrig

International Advisory Committee:
Nestor Armesto (Univ. de Santiago de Compostela, Spain)
Elke Aschenauer (BNL, USA)
Daniel Boer (University of Groningen, Netherlands)
Marco Contalbrigo (INFN Ferrara, Italy)
Markus Diehl (DESY, Germany)
Rolf Ent (JLab, USA)
Max Klein (University of Liverpool, UK)
Andrzej Sandacz (National Centre for Nuclear Research, Poland)
Marco Stratmann (University of Tübingen, Germany)
Lech Szymanowski (National Centre for Nuclear Research, Poland)
Tony Thomas (University of Adelaide, Australia)
Thomas Ullrich (BNL, USA)
Raju Venugopalan (BNL, USA)

Topics:
■ Structure of hadrons: (nuclear) parton distribution functions (PDFs, nPDFs), transverse momentum dependent (TMDs) and generalized parton distributions (GPDs), Distribution Amplitudes (DAs), Double Distributions (DDs),
■ QCD at high parton densities and small- x : saturation, evolution, Color Glass Condensate
■ Fragmentation functions and Jet properties
■ Complementarity and connections of EIC physics with p+p, p+A and A+A collisions: high p_T processes, diffraction, multi-parton interactions, quark-gluon plasma and colored probes in hot nuclear matter,
■ Physics beyond the Standard Model and connections to other areas in physics,
■ Future DIS facilities: accelerator and detector developments.

Satellite workshop during POETIC 8

The poster for MCEGs 2019 features a background image of a city skyline at sunset with a large, colorful, abstract circular pattern. The text is white and orange.

February 20-22, 2019
DESY Hamburg, Germany

EIC User Group and MCnet present
MCEGs
for future ep and eA facilities

PROGRAM
Updates to general-purpose MCEG for ep / eA
Status of NLO simulations for ep/eA
GPDs and TMDs in MCEGs
QED+QCD effects in ep/eA simulations

ORGANIZERS
Elke-Caroline Aschenauer (BNL) Simon Plätzer (University of Vienna)
Andrea Bressan (INFN Trieste) Stefan Prestel (Lund University)
Markus Diefenthaler (JLAB)
Hannes Jung (DESY)

www.desy.de/mceg2019

The poster for MCEGs 2019 features a background image of a city skyline at sunset with a large, colorful, abstract circular pattern. The text is white and orange.

November 20-22, 2019
Erwin-Schrödinger Institute
Vienna, Austria

EIC User Group and MCnet present
MCEGs
for future ep and eA facilities

PROGRAM
MCEGs for eA, including light and heavy ions
Validation of HERA data
MCEGs for TMDs

ORGANIZERS
Elke-Caroline Aschenauer (BNL) Simon Plätzer (Vienna)
Andrea Bressan (Trieste) Stefan Prestel (Lund)
Markus Diefenthaler (JLAB)
Hannes Jung (DESY)

<https://indico.cern.ch/event/845653/>

Organized by Elke-Caroline Aschenauer (BNL), Andrea Bressan (Trieste), Markus Diefenthaler (JLab), Hannes Jung (DESY), Simon Plätzer (Vienna), Stefan Prestel (LUND)

Summary from MCEG workshop series

MCEG for ep On a good path, but still a lot of work ahead.

- **General-purpose MCEGs**, HERWIG, PYTHIA, and SHERPA, will be significantly improved w.r.t. MCEGs at HERA time:
- Comparisons with HERA data and QCD predictions critical:
 - To learn where physics models need to be improved,
 - To complement MC standard tunes with first DIS/HERA tune.
- The existing general-purpose MCEG should be able to simulate NC and CC unpolarized observables also for eA. A precise treatment of the nucleus and, e.g., its breakup is needed.
- First parton showers and hadronization models for ep with spin effects, but far more work needed for polarized ep / eA simulations.
- Need to clarify the details about merging higher QED+QCD effects (in particular for eA).

MCEG for eA Less clear situation about theory and MCEG.

- **Pioneering projects**, e.g., BeAGLE, spectator tagging in ed, Sartre.
- **Active development**, e.g., eA adaptation of JETSCAPE, Mueller dipole formalism in Pythia8 (ala DIPSY).


Introducing modern general-purpose MCEGs and Rivet

EICUG Software Working Group

EIC SOFTWARE TUTORIALS

01/09	Introductory Tutorials
01/29	Detector Full Simulation
02/06	Detector Full Simulation
04/21	Advanced Fast Simulations
08/17-18,20,26	Monte Carlo Event Generators

FIND MORE TUTORIALS ON YOUTUBE:



EICUG EIC User Group
119 subscribers

HOME VIDEOS PLAYLISTS CHANNELS ABOUT

Uploads ▶ PLAY ALL

EIC Software Tutorial: Pythia 8
1.4K views • 1 year ago

EIC Software Tutorial: Herwig 7 Overview
223 views • 1 year ago

EIC Software Tutorial: MC-Data Comparisons in Rivet
196 views • 1 year ago

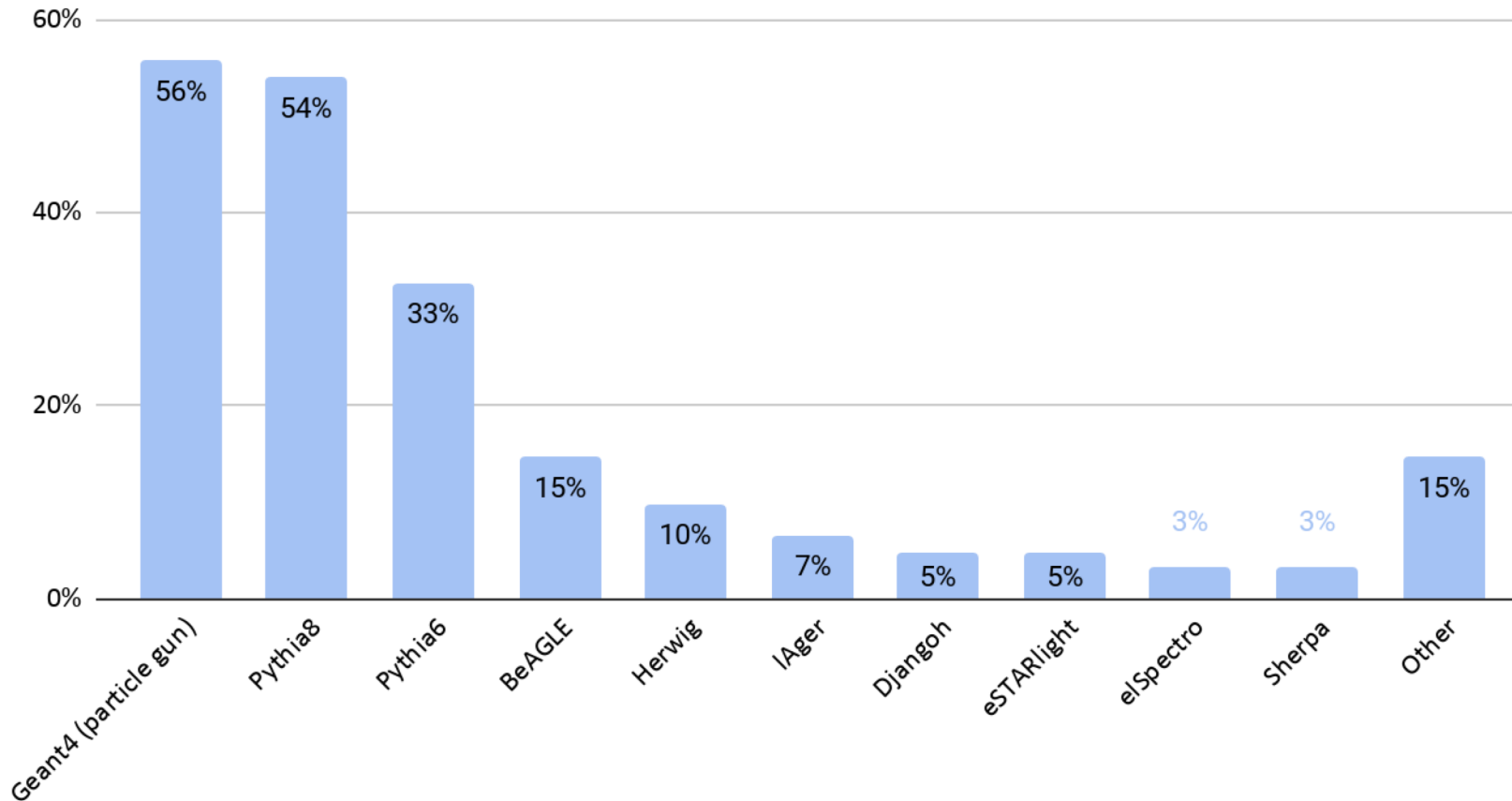
EIC Software Tutorial: Sherpa
169 views • 1 year ago

Excellent feedback on online tutorials and their recordings.

MCEGs used for Yellow Report

Source [State of Software Survey](#)

N = 61, average number of selected options = 2.0



Other (N = 9): personal computer codes (N = 2), ACT, CLASDIS, ComptonRad, GRAPE-DILEPTON, MADX, MILOU, OPERA, RAYTRACE, Sartre, Topeg, ZGOUBI

Starting with MCEG validation using Rivet

→ Detailed presentation by **Vaibhavi Gawas**

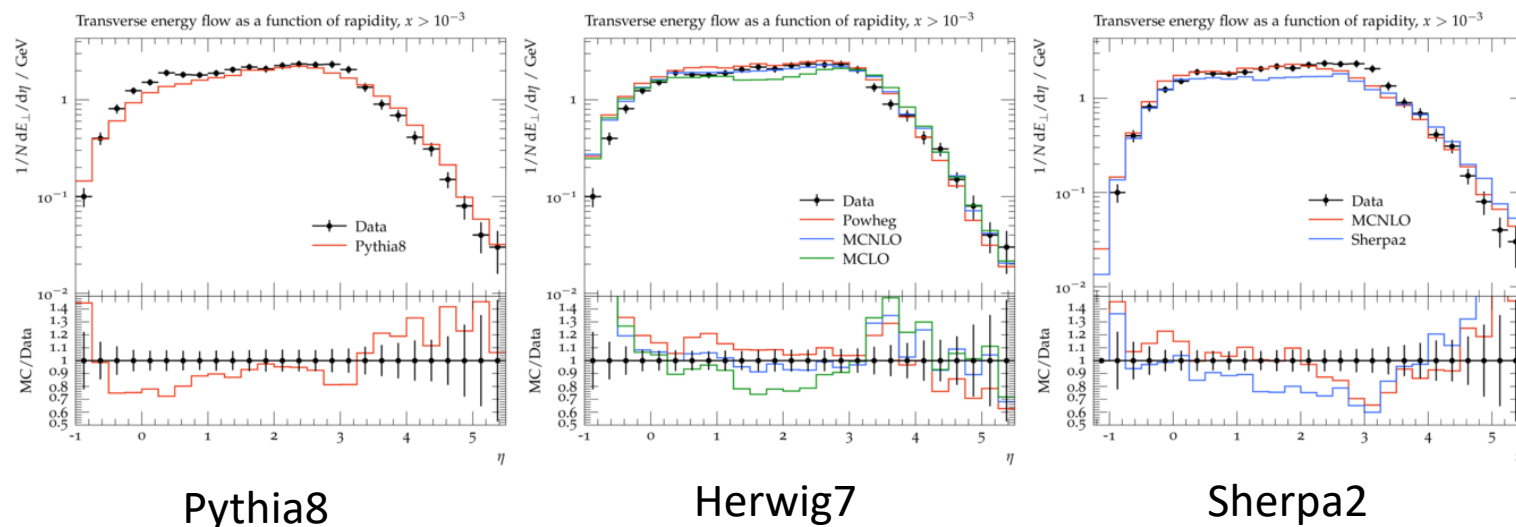
MCEG R&D requires *easy* access to *data*:

- data := analysis description + data points

HEP existing workflow using Rivet.

Ongoing activity with EIC-India and MCnet:

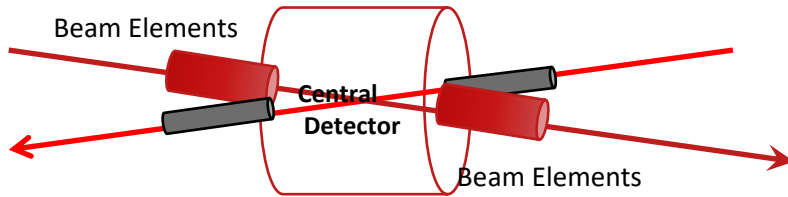
- Comparison to published results using RIVET and understand differences.
- **Provide initial findings and results in publication (work in progress)::**
 - Overview of where we stand in understanding HERA data with current physics and models implement in MCEGs.



Machine-Detector interface (MDI)

Integrated interaction region and detector design to optimize physics reach

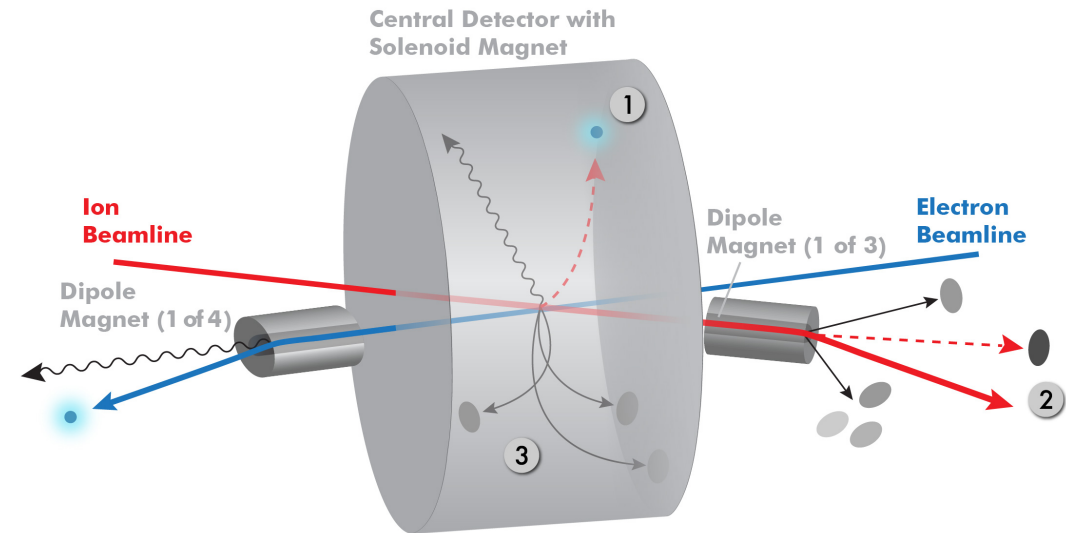
The aim is to get **~100% acceptance** for all final state particles, and measure them with good resolution.



Experimental challenges:

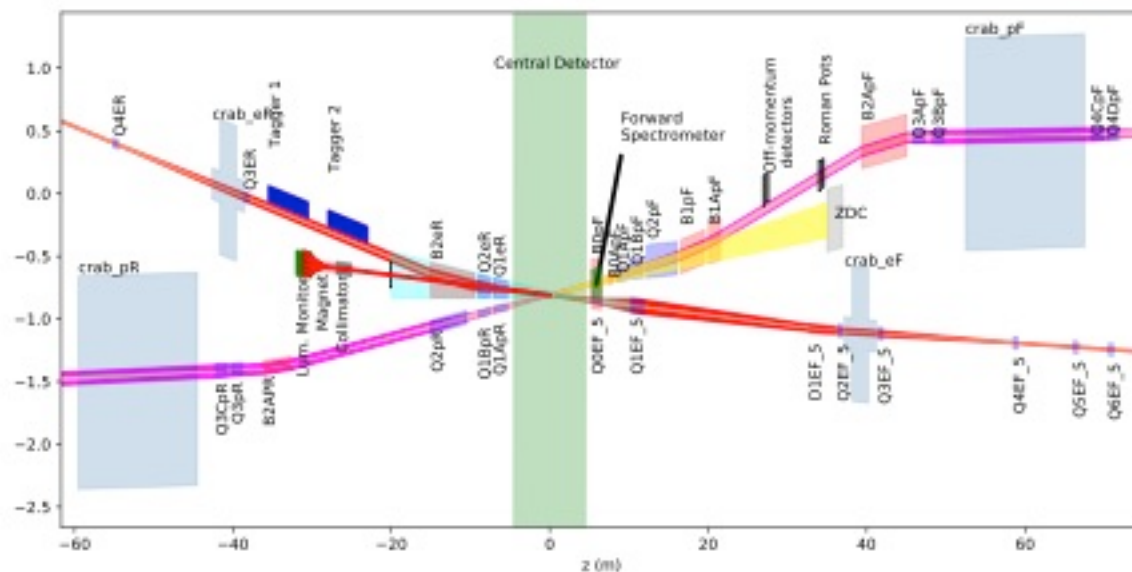
- beam elements limit forward acceptance
- central Solenoid not effective for forward

Possible to get ~100% acceptance for the whole event.



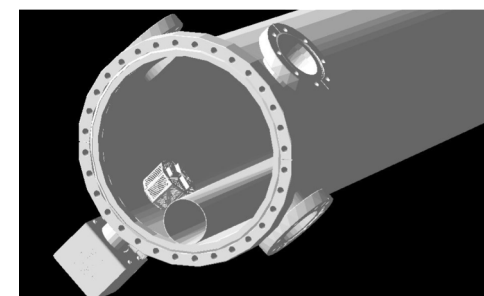
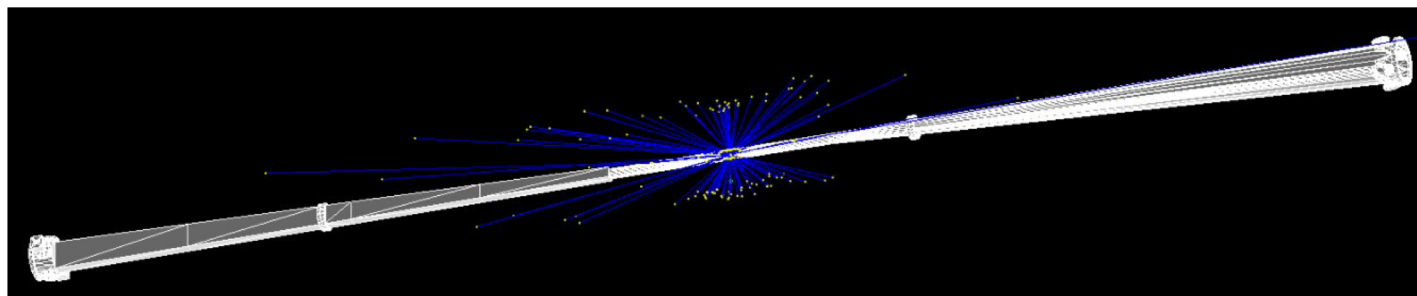
MDI in Simulations

IR Layout



Unprecedented integration of IR and detector (shown here for IP6).

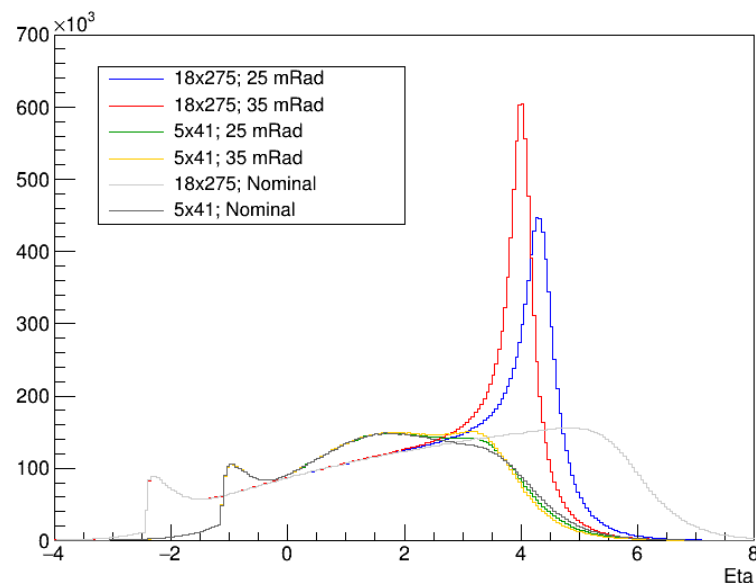
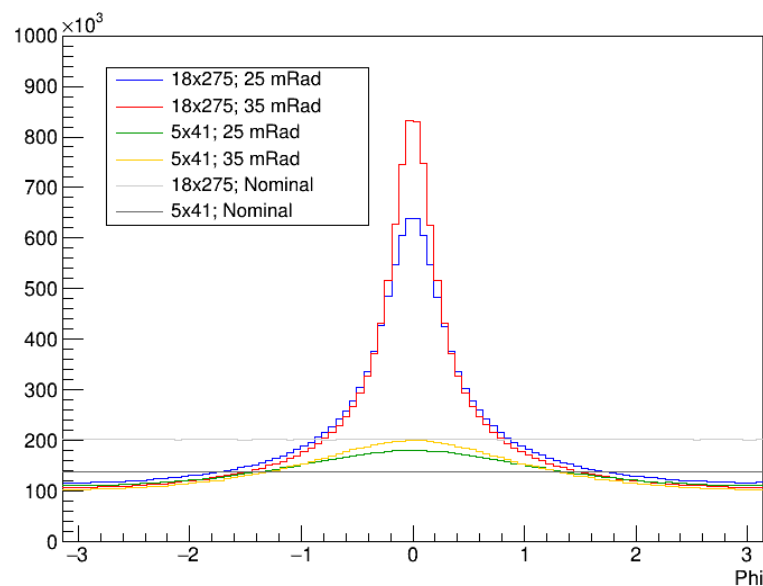
CAD Interface (accelerator elements and service structures)



EIC Project

Simulation based (in part) on CAD files provided by EIC project engineering teams, rather than a bottoms-up reliance on constructive solid geometry (Screenshots from **eAST**)

Accelerator and Beam Conditions Critical for EIC Simulations



- Accelerator and beam effects that influence EIC measurements

- Beam crossing angle,
- Crabbing rotation,
- Beam energy spread,
- Angular beam divergence,
- Beam vertex spread.

- Note for EIC Community <https://eic.github.io/resources/simulations.html>

- Profound consequences on measurement capabilities of the EIC and layout of the detectors,
- How to integrate these effects in EIC simulations.
- **Authors** J. Adam, E.-C.Aschenauer, M. Diefenthaler, Y. Furletova, J. Huang, A. Jentsch, B. Page.

Beyond that Include beam background estimates in simulations.

Discussion of Event Generation and Simulation Needs

Simulation of physics processes

Monte Carlo Event Generators

Simulation of detector responses

Fast simulations

Full simulations

Analysis of simulated data

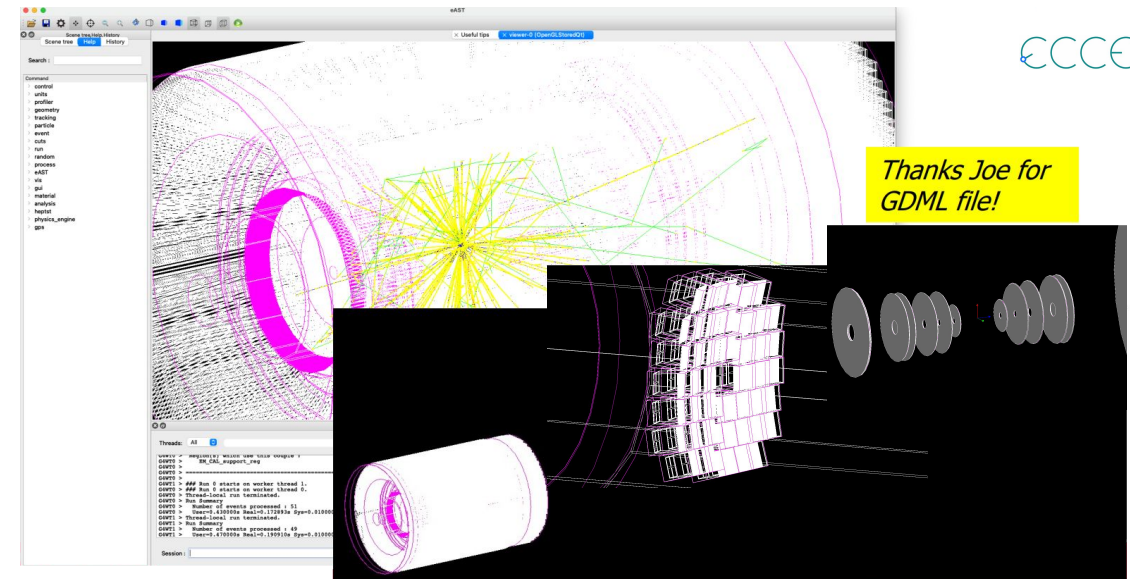
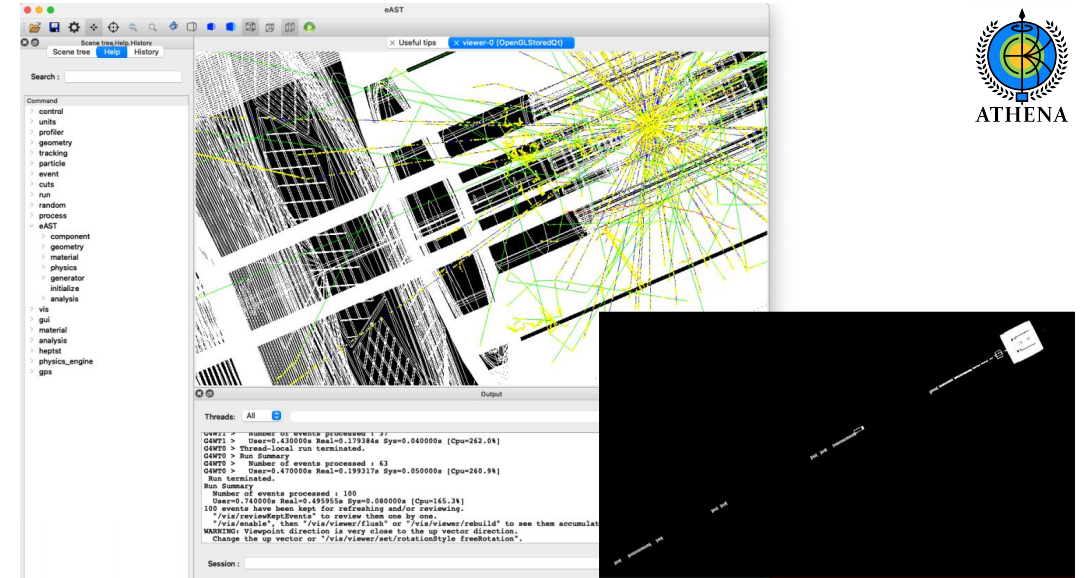
Reconstruction

Detector simulations and Geant4

(Screenshots from eAST)

EIC

- Detector (and physics) simulations rely on **Geant4**, the (!) detector simulation toolkit for HEP and NP:
 - Detector full simulations for ATHENA and ECCE detector concepts based on **Geant4**.
 - As GeantV comes up at times:
 - Project concluded: no performance gain from the vectorization of the individual software components,
 - Modular software packages such as VecGeom integrated into **Geant4**.
- Energy range is different from LHC,
- validation, tuning and extension including test beam studies required.
- Ongoing collaboration with international Geant4 collaboration, including Technical Forum on NP/EIC.



Towards a Next-Generation Simulations



ELECTRON ION COLLIDER USER GROUP STATE OF SOFTWARE SURVEY

Survey from February 16 – 23, 2021.

[report](#)

There are **too many generators and simulation tools** used at the moment.

Unify the Simulation Effort

- The SWG is preparing to launch a **common effort on next-generation simulations**:
 - building on the work done in the existing simulations,
 - unify the software community behind one common effort,
 - a requirement for the common framework is that it integrate the existing detector simulations in a modular way.



Detector Simulation

- **comprehensive, centrally maintained application**
- **based on Geant4 for fast and full simulations**
- **with library of potential detector options**

Requirements

- ability to **reuse existing simulation work**
- ease of **switching detector options**
- ease of switching between **detailed and coarse** detector descriptions
- **ease of leveraging new and rapidly evolving technologies,**
 - AI/ML to accelerate simulations
 - computing hardware, e.g., heterogeneous architectures
 - AI/ML is the best near term prospect for using LCF/Exascale effectively

Project Leader

- Makoto Asai, Geant4 project leader and deep technical expert for >20yrs.



The role of AI/ML in simulations

Lesson learned High-precision QCD measurements require high-precision simulations

Statistical accuracy for precise hypothesis testing

- up to trillion of simulated events required (HL-LHC)
- often computationally intensive, in particular calorimeter simulations

Common alternatives

- fast simulations with computationally efficient approximations, e.g., parameterizations or look-up tables
- **still** insufficient accuracy for high-precision measurements

Promising alternatives

- fast generative models, e.g., GANs or VAEs
- AI driven design, e.g., Bayesian optimization

Eur. Phys. J. A (2021) 57:100
<https://doi.org/10.1140/epja/s10050-020-00290-x>

THE EUROPEAN
PHYSICAL JOURNAL A



Review

A.I. for nuclear physics

Paulo Bedaque¹, Amber Boehnlein^{2,a}, Mario Cromaz³, Markus Diefenthaler², Latifa Elouadrhiri², Tanja Horn⁴, Michelle Kuchera⁵, David Lawrence², Dean Lee⁶, Steven Lidia⁶, Robert McKeown², Wally Melnitchouk², Witold Nazarewicz⁶, Kostas Orginos^{2,7}, Yves Roblin², Michael Scott Smith⁸, Malachi Schram⁹, Xin-Nian Wang³

¹ University of Maryland, College Park, MD, USA

² Thomas Jefferson National Accelerator Facility, Newport News, VA, USA

³ Lawrence Berkeley National Laboratory, Berkeley, CA, USA

⁴ Catholic University, Washington D.C., USA

⁵ Davidson College, Davidson, NC, USA

⁶ Michigan State University, East Lansing, MI, USA

⁷ College of William & Mary, Williamsburg, VA, USA

⁸ Oak Ridge National Laboratory, Oak Ridge, TN, USA

⁹ Pacific Northwest National Laboratory, Richland, WA, USA

Received: 9 September 2020 / Accepted: 7 October 2020 / Published online: 22 March 2021

© Jefferson Science Associates, LLC, under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Communicated by Ulf Meissner

Abstract This report is an outcome of the workshop AI for Nuclear Physics held at Thomas Jefferson National Accelerator Facility on March 4–6, 2020

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177. Participation of students and early career professionals was supported by NSF, Division of Physics, under the Grant ‘Artificial Intelligence (AI) Workshop in Nuclear Physics,’ Award Number 2017170. Support for the Hackathon was provided by the University of Virginia School of Data Sciences and by Amazon Web Services.

This report is an outcome of the workshop AI for Nuclear Physics held at Thomas Jefferson National Accelerator Facility on March 4–6, 2020. The workshop brought together 184 scientists to explore opportunities for Nuclear Physics in the area of Artificial Intelligence. The workshop consisted of plenary talks, as well as six working groups.

^a e-mail: amber@jlab.org (corresponding author)

Contents

1 Executive summary	2
2 Priority research directions	3
2.1 Future prospects	3
2.2 Community identified needs and commonalities	5
3 Workshop overview	6
4 Summary of workshop sessions	6
4.1 Lattice QCD and other quantum field theories	6
4.1.1 Case studies and future prospects	7
4.1.2 Enabling discoveries/what is needed	7
4.2 Low-energy nuclear theory	7
4.2.1 Current status	7
4.2.2 Case studies and future prospects	8
4.2.3 Enabling discoveries/what is needed	9
4.3 Accelerator science and operations	10
4.3.1 Accelerator and material design optimization	10
4.3.2 Provenance and prognostication for accelerator sub-systems	10
4.3.3 Dynamic optimization of real time operation controls	11
4.3.4 Summary and Final Thoughts	11
4.4 Experimental methods	12
4.4.1 Current status	12
4.4.2 Case Studies and Future Prospects	12
4.4.3 Enabling discoveries/what is needed	13
4.5 Event generation and simulation	13
4.5.1 Current status	13
4.5.2 Case studies and future prospects	14

Springer

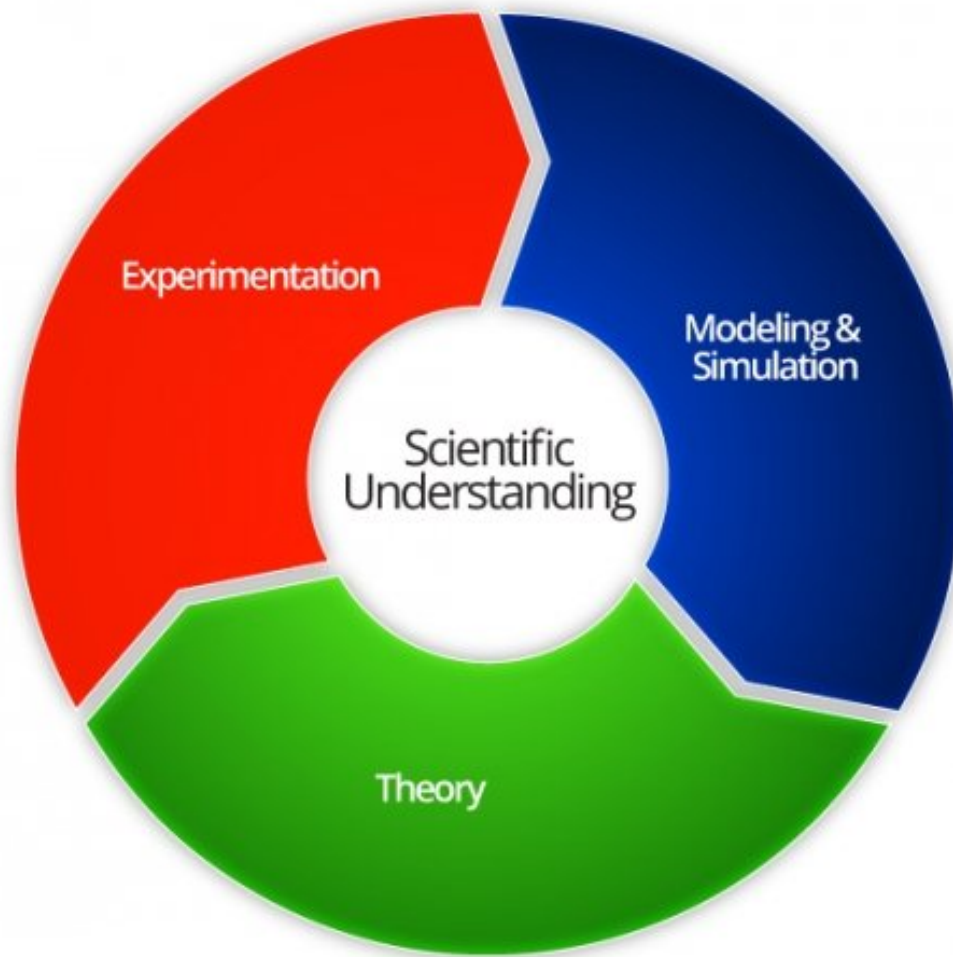
Bringing experts in various domains of QCD theory and experiment together

*“It will be **joint progress of theory and experiment** that moves us forward, not in one side alone”*
Donald Geesaman (ANL, former NSAC Chair) on “[Nuclear Physics in a Decade](#)”

Open Questions

- **How will we compare theory and experiment at the EIC?**
 - Will we unfold our experimental measurements to Born level and compare them to theoretical calculations at Born level? This has been done at HERA.
 - Will we fold the theoretical calculations with radiative and detector effects and compare them at detector level? This has been done for NP and being developed further, e.g., [[arXiv:2108.13371](#)]. This is also being [discussed](#) for LHC.
 - This has profound consequences on the reproducibility and reinterpretability of our measurements and the requirements on our analysis software and workflows.
- **How will we evaluate PDFs and FFs at the EIC?**
 - Can we have a “LHAPDF” type of interface for PDF and FFs, including nuclear, spin-dependent, transverse-momentum dependent effects? We need this for LO, NLO, and beyond.
 - Can we use ML to parametrize all of this information, including uncertainties, in a computationally efficient way?
- **Do we understand the limitations of measurements at the EIC?**
 - We are better and better understanding the measurement capabilities of the EIC detectors.
 - What about the theoretical limitations of the measurements? How can we address them?

Towards high-precision simulations for the EIC



Next steps for MC4EIC

- Discuss **requirements for MCEGs** and related computations tools to simulate the collision of highly-polarized electrons and highly-polarized light ions and unpolarized heavy ions.
- Develop a **roadmap for MCEG developments for the EIC**, including HEP-NP funding.
- **Priorities for the next years** could be:
 - Training of the EIC community, e.g., via online tutorials.
 - Validation of existing MCEGs using Rivet. Build automated workflows (CI/CD).
 - Development of a DIS tune.
 - Merging of higher order QED and QCD effects.
 - Interface between MCEGs and Geant4 based on HepMC3.
 - Roadmap for spin-dependent parton showers.
 - Roadmap for spin-dependent hadronization models.
 - Guidance on how to compare measurements at the EIC with theory.

Summary

mdiefent@jlab.org



- **Simulations** essential for design of experiments, data analysis, and verification of measurements.
- **Simulations** for the **EIC**, i.e. MCEGs and fast and full detector simulations for the EIC, require **R&D**. We miss core capabilities and we need to work towards accuracy and precision.
- **Simulation R&D** is most efficiently done in common projects and in collaboration with other fields, in particular HEP.
- Many opportunities for AI/ML to complement and improve **simulations**. While AI/ML approaches will substitute parts of our **simulation workflows**, they will not replace core tools, e.g., general-purpose MCEGs or Geant4.